

# Structural Fluctuation of Protein in Water around Its Native State: A New Statistical Mechanics Formulation

Bongsoo Kim<sup>1</sup> and Fumio Hirata<sup>2</sup>

<sup>1</sup> Department of Physics and Institute for Soft and Bio Matter Science,

Changwon National University, Changwon 641-773, Korea

<sup>2</sup> College of Life Sciences, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan

(Dated: October 31, 2012)

A new statistical mechanics formulation of characterizing the structural fluctuation of protein correlated with that of water is presented based on the generalized Langevin equation and the 3D-RISM/RISM theory of molecular liquids. The displacement vector of atom positions and their conjugated momentum, are chosen for the dynamic variables for protein, while the density fields of atoms and their momentum fields are chosen for water. Projection of other degrees of freedom onto those dynamic variables using the standard projection operator method produces essentially two equations which describe the time evolution of fluctuation concerning the density field of solvent and the conformation of protein around an equilibrium state, which are coupled with each other. The equation concerning the protein dynamics is formally akin to that of the coupled Langevin oscillators, and is a generalization of the latter, to atomic level. The most intriguing feature of the new equation is that it contains the variance-covariance matrix as the "Hessian" term describing the "force" restoring an equilibrium conformation, which is the second moment of the fluctuation of atom positions. The "Hessian" matrix is naturally identified as the second derivative of the free energy surface around the equilibrium. A method to evaluate the Hessian matrix based on the 3D-RISM/RISM theory is proposed. Proposed also is an application of the present formulation to the molecular recognition, in which the conformational fluctuation of protein around its native state becomes an important factor as exemplified by so called "induced fitting".

## I. INTRODUCTION

Structural fluctuation of protein around its native state plays essential roles in a variety of processes in which the biomolecule performs its intrinsic function [1]. For example, so called "gating" mechanisms of ion channels are regulated by the structural fluctuation of amino-acid residues consisting the gate region of the channel. Molecular recognition such as the formation of an enzyme-substrate complex in an enzymatic reaction is controlled often by structural fluctuation of protein. Few typi-

cal examples of structural fluctuations around a native conformation of protein, related to function, are "breathing" [2], "hinge-bending" [3], and "arm-rotating" motions [4]. Those motions are *collective* in nature involving many atoms moving in the same direction. Those structural fluctuation associated with protein functions, whether it's large or small, stays within its native conformation, and does not induce global conformational change such as denaturing, with few exceptions exemplified by *intrinsically disordered protein* [5].

In actual biological processes, solvent plays vital roles both in the equilibrium and in fluctuation of protein [6]. It may not be necessary to spend many words for emphasizing the crucial role played by solvent for stabilizing or destabilizing native structure of protein, such as the hydrophobic interaction and hydrogen bonds. Here, let us consider roles played by water in fluctuation of protein around its native conformation, associated with recognition of a ligand by protein. The process is primarily a thermodynamic process, governed by the free energy difference between the two states before and after the recognition. It is obvious that water plays crucial role in the thermodynamics, since the equilibrium structures are determined by the free energies including the excess chemical potential or the solvation free energy of water. However, it is not the only role of water in the process. Water actually regulates the kinetic pathway of the process as well by controlling the structural fluctuation of amino-acid residues consisting the active site. An example of such processes is a mouth-like motion of amino-acid residues. The open-and-close motion of the mouth is driven not only by the direct force acting among atoms in protein, but by that originated from the solvent induced force which is in turn caused by the fluctuation in the solvation free energy, or the non-equilibrium free energy. In an actual biomolecular process, such conformational change around the native state is induced often by some perturbation upon amino-acid residues around the active site, for example, binding of a ligand. However, response to the perturbation should be linear, because the protein recovers its native conformation upon removing the perturbation [7].

It is not surprising that considerable efforts have been devoted to clarify the conformational fluctuation of protein theoretically, which has started at the end of the last century based on the molecular mechanics or dynamics. One of earliest attempts was to relate the structural fluctuation to the normal mode of protein [8]. Those works have demonstrated the importance of the collective mode in the fluctuation. However, those efforts have not provided a realistic physical insight into the dynamics of actual biological processes, since they are concerned with a protein in "vacuum", which obviously cannot describe the fluctuation correlated with that of solvent. The principal component analysis involving diagonalization of the variance-covariance matrix of conformational fluctuation, extracted

from the molecular dynamics trajectory of a protein in water, has revealed some important aspects of the conjugated fluctuation between a biomolecule and water [9]. The lowest frequency mode of fluctuation around a native conformation exhibits an activated transition from a minimum to another minimum in the conformational space, akin to the jump diffusion model of liquids. However, the procedure cannot be extended readily to that associated with such a process as ligand binding, because the process is concerned with sampling of large configuration space involving both protein and solvent. It becomes formidable especially when the solvent consists of several chemical components such as the electrolyte solution.

In the present work, we propose a new first-principle approach to treat structural fluctuation of protein conjugated with that of solvent, based on the two theoretical frameworks in the statistical mechanics of liquids, or, the 3D-RISM/RISM (Reference Interaction Site Model) theory and the generalized Langevin equation [10]. The 3D-RISM/RISM theory [11] has proven itself to be capable of predicting the molecular recognition of ligand by protein which has a rigid structure [12]. The generalized Langevin equation should be able to describe the fluctuation of a system consisting of protein and solvent around its equilibrium state. Therefore, it is reasonable to expect that the two theories combined together will produce a method which can describe the molecular recognition process by protein, whose structure is fluctuating. An etude of such a theory has been already published by us [13] where a much more simplified model, a chain of identical particles in solvent consisting of spherical molecules, was considered. The key idea there lies in the choice of dynamic variables. We have chosen four quantities to form a vector in the phase space: the displacement of atom positions in protein from their equilibrium coordinates, the conjugated momentum of those atoms, the fluctuation of the density field of solvent molecules, and their conjugated momentum field or flux. A standard treatment of the dynamic variables due to the projection operator method [10, 14] gave rise to four equations with respect to the time evolution of those quantities, two for solute and two for solvent, which interplay with each other. Most important observation in the results is that the equation of motion concerning the solute dynamics includes the variance-covariance matrix regarding the conformational fluctuation of solute as a "Hessian" or a "force constant" of the "oscillation" or fluctuation. Here, we generalize the theory developed in the preceding paper [13] substantially in order to be able to treat a realistic protein in a realistic solvent such as water.

## II. PROJECTION OPERATOR METHOD: SUMMARY

Since the projection operator method is well-known [15–17], we here only summarize the general results of the method. It gives the time evolution equation of a dynamic variable  $\mathbf{A}(t)$  which is a function of microscopic variables. Its microscopic time evolution is governed by the Liouville operator  $i\mathcal{L}$  whose expression will be given in the next section:

$$\frac{d\mathbf{A}(t)}{dt} \equiv i\mathcal{L}\mathbf{A}(t) \equiv \{A, \mathcal{H}\}_{PB}(t) \quad (\text{II.1})$$

where  $\{a, b\}_{PB}$  is the Poisson bracket, and  $\mathcal{H}$  is the Hamiltonian of the system. The formal solution of (II.1) is given by

$$\mathbf{A}(t) = e^{i\mathcal{L}t} \mathbf{A}(0) \equiv e^{i\mathcal{L}t} \mathbf{A} \quad (\text{II.2})$$

Now the projection operator  $\mathcal{P}$  is defined as

$$\mathcal{P}(\cdots) \equiv (\mathbf{A}, \cdots)(\mathbf{A}, \mathbf{A})^{-1} \mathbf{A} \quad (\text{II.3})$$

The inner product  $(\mathbf{a}, \mathbf{b})$  denotes an average of the canonical distribution  $\exp(-\mathcal{H}/k_B T)$ :

$$(\mathbf{a}, \mathbf{b}) \equiv \langle \mathbf{a}^* \mathbf{b} \rangle = \mathcal{Z}^{-1} \int d\Gamma \mathbf{a}^* \mathbf{b} \exp(-\mathcal{H}(\Gamma)/k_B T) \quad (\text{II.4})$$

where  $\Gamma$  denotes all microscopic degrees of freedom in the system. The operator projects out only the 'component' of  $\mathbf{A}$  from the object  $(\cdots)$ . Then obviously  $\mathcal{P}\mathbf{A} = \mathbf{A}$  holds. It also has the idempotent property  $\mathcal{P}^2 = \mathcal{P}$ .

After projecting  $\mathbf{A}$ -component out of the microscopic degrees of freedom, the exact time evolution equation for  $\mathbf{A}(t)$  is given by

$$\frac{d\mathbf{A}(t)}{dt} = i\boldsymbol{\Omega} \cdot \mathbf{A}(t) - \int_0^t ds \mathbf{K}(t-s) \cdot \mathbf{A}(s) + \mathbf{f}(t) \quad (\text{II.5})$$

Here the frequency matrix  $i\boldsymbol{\Omega}$ , the memory matrix  $\mathbf{K}(t)$ , and the fluctuating force vector  $\mathbf{f}(t)$  are given by

$$\begin{aligned} i\boldsymbol{\Omega} &= (\mathbf{A}, \dot{\mathbf{A}}) \cdot (\mathbf{A}, \mathbf{A})^{-1}, \\ \mathbf{K}(t) &= (\mathbf{f}, \mathbf{f}(t)) \cdot (\mathbf{A}, \mathbf{A})^{-1}, \\ \mathbf{f}(t) &= \exp(t(1 - \mathcal{P})i\mathcal{L})(1 - \mathcal{P})\dot{\mathbf{A}} \end{aligned} \quad (\text{II.6})$$

One can show easily that the fluctuating force  $\mathbf{f}(t)$  does not have  $\mathbf{A}$ -component, i.e.,  $(\mathbf{A}, \mathbf{f}(t)) = 0$ . Using this feature and the linearity of the equation, we immediately obtain the following dynamic

equation for the auto-correlation function of  $\mathbf{A}(t)$ ,  $\mathbf{C}(t)$

$$\frac{d\mathbf{C}(t)}{dt} = i\boldsymbol{\Omega} \cdot \mathbf{C}(t) - \int_0^t ds \mathbf{K}(t-s) \cdot \mathbf{C}(s) \quad (\text{II.7})$$

### III. GENERALIZED LANGEVIN EQUATIONS FOR A SOLUTE-SOLVENT SYSTEM

Our main concern here is a protein-water system at infinite dilutions. However, the formulation is completely general for any solute-solvent system at infinite dilution. So, in the formulation, we consider a general solute-solvent system. In particular, we consider a solute molecule consisting of  $N_u$  atoms immersed in solvent consisting of  $N$  molecules, each having  $n$  atoms. The Hamiltonian of the solute-solvent system is then given by

$$\begin{aligned} \mathcal{H} &\equiv \mathcal{H}_0 + \mathcal{H}_1 + \mathcal{H}_2, \\ \mathcal{H}_0 &= \sum_{i=1}^N \sum_{a=1}^n \left[ \frac{\mathbf{p}_i^a \cdot \mathbf{p}_i^a}{2m_a} + \sum_{j \neq i} \sum_{b \neq a} U_0(|\mathbf{r}_i^a - \mathbf{r}_j^b|) \right] \quad (\text{solvent}) \\ \mathcal{H}_1 &= \sum_{\alpha=1}^{N_u} \left[ \frac{\mathbf{P}_{\alpha} \cdot \mathbf{P}_{\alpha}}{2M_{\alpha}} + \sum_{\beta \neq \alpha} U_1(|\mathbf{R}_{\alpha} - \mathbf{R}_{\beta}|) \right] \quad (\text{solute}) \\ \mathcal{H}_2 &= \sum_{\alpha=1}^{N_u} \sum_{i=1}^N \sum_{a=1}^n U_{int}(|\mathbf{R}_{\alpha} - \mathbf{r}_i^a|) \quad (\text{solute-solvent}) \end{aligned} \quad (\text{III.1})$$

where  $M_{\alpha}$  denotes the mass of the  $\alpha$ th atom in the solute particle, and  $m_a$  the mass of  $a$ th atom in a solvent molecule. The Hamiltonian of the solvent is denoted by  $\mathcal{H}_0$  where  $\mathbf{r}_i^a$  and  $\mathbf{p}_i^a$  are respectively the position and momentum of  $a$ th atom in the  $i$ th molecule of the solvent, and  $U_0(r_{ij}^{ab})$  ( $r_{ij}^{ab} \equiv |\mathbf{r}_i^a - \mathbf{r}_j^b|$ ) is the pair potential energy between them.  $\mathcal{H}_1$  is the Hamiltonian of the  $N_u$  solute atoms, and  $\mathbf{R}_{\alpha}$  and  $\mathbf{P}_{\alpha}$  are the position and momentum of the  $\alpha$ th solute atom (we preserve the Greek indices for denoting the solute atoms), and  $U_{int}(|\mathbf{R}_{\alpha} - \mathbf{r}_i^a|)$  is the interaction potential energy between the  $\alpha$ th solute atom and the  $a$ th atom of the  $i$ th molecule in the solvent.

The associated Liouville operator  $i\mathcal{L}$  is given by

$$\begin{aligned} i\mathcal{L} &\equiv i\mathcal{L}_0 + i\mathcal{L}_1, \\ i\mathcal{L}_0 &\equiv \sum_{i=1}^N \sum_{a=1}^n \left[ \frac{1}{m_a} \mathbf{p}_i^a \cdot \frac{\partial}{\partial \mathbf{r}_i^a} - \sum_{j \neq i} \sum_{b \neq a} \frac{\partial U_0(r_{ij}^{ab})}{\partial \mathbf{r}_i^a} \cdot \frac{\partial}{\partial \mathbf{p}_i^a} - \sum_{\alpha=1}^{N_u} \frac{\partial U_{int}(|\mathbf{R}_{\alpha} - \mathbf{r}_i^a|)}{\partial \mathbf{r}_i^a} \cdot \frac{\partial}{\partial \mathbf{p}_i^a} \right] \\ i\mathcal{L}_1 &\equiv \sum_{\alpha=1}^{N_u} \left[ \frac{\mathbf{P}_{\alpha}}{M_{\alpha}} \cdot \frac{\partial}{\partial \mathbf{R}_{\alpha}} + \mathbf{F}_{\alpha} \cdot \frac{\partial}{\partial \mathbf{P}_{\alpha}} \right] \end{aligned} \quad (\text{III.2})$$

where  $\mathbf{F}_{\alpha} \equiv \mathbf{F}_{\alpha}^{(u)} + \mathbf{F}_{\alpha}^{(v)}$ , and  $\mathbf{F}_{\alpha}^{(u)}$  is the force exerted on the  $\alpha$ th solute atom by the other solute atoms,  $\mathbf{F}_{\alpha}^{(v)}$  the force exerted on the same solute atom by the solvent molecules. Their explicit expressions

are given by

$$\mathbf{F}_\alpha^{(u)} = - \sum_{\beta \neq \alpha} \frac{\partial U_1(R_{\alpha\beta})}{\partial \mathbf{R}_\alpha}, \quad \mathbf{F}_\alpha^{(v)} = - \sum_{i=1}^N \sum_{a=1}^n \frac{\partial U_{int}(|\mathbf{R}_\alpha - \mathbf{r}_i^a|)}{\partial \mathbf{R}_\alpha} \quad (\text{III.3})$$

Our dynamic variable  $\mathbf{A}(t)$  is chosen to be

$$\mathbf{A}(t) = \begin{pmatrix} \Delta \mathbf{R}_\alpha(t) \\ \mathbf{P}_\alpha(t) \\ \delta \rho_{\mathbf{k}}^a(t) \\ \mathbf{J}_{\mathbf{k}}^a(t) \end{pmatrix} \quad (\text{III.4})$$

Here  $\Delta \mathbf{R}_\alpha(t)$  is the displacement of the position vector  $\mathbf{R}_\alpha$  of the  $\alpha$ -th solute atom from its equilibrium value. And  $\delta \rho_{\mathbf{k}}^a(t)$  is the Fourier component of the density fluctuation  $\delta \rho^a(\mathbf{r}, t) \equiv \rho^a(\mathbf{r}, t) - \rho_0^a$  ( $\rho_0^a$  is the average number density of the  $a$ th atom) of the solvent liquid:

$$\begin{aligned} \delta \rho^a(\mathbf{r}, t) &\equiv \sum_i \delta(\mathbf{r} - \mathbf{r}_i^a(t)) - \rho_0^a, \\ \delta \rho_{\mathbf{k}}^a(t) &= \int d\mathbf{r} e^{i\mathbf{k} \cdot \mathbf{r}} \delta \rho^a(\mathbf{r}, t) = \sum_i e^{i\mathbf{k} \cdot \mathbf{r}_i^a(t)} - (2\pi)^3 \rho_0^a \delta(\mathbf{k}) \end{aligned} \quad (\text{III.5})$$

Likewise,  $\mathbf{J}_{\mathbf{k}}^a(t)$  is the Fourier component of the current of the  $a$ th atom in the solvent liquid:

$$\begin{aligned} \dot{\rho}_{\mathbf{k}}^a(t) &= \sum_i i\mathbf{k} \cdot \frac{\mathbf{p}_i^a(t)}{m_a} e^{i\mathbf{k} \cdot \mathbf{r}_i^a(t)} \equiv i\mathbf{k} \cdot \mathbf{J}_{\mathbf{k}}^a(t) \\ \mathbf{J}_{\mathbf{k}}^a(t) &= \sum_i \frac{\mathbf{p}_i^a(t)}{m_a} e^{i\mathbf{k} \cdot \mathbf{r}_i^a(t)} \end{aligned} \quad (\text{III.6})$$

Balucani and Zoppi [15] have worked out the special case where only the momentum of a solute particle was chosen as a dynamic variable.

We now proceed to obtain the specific expressions of the eqs. (II.5) or (II.6). First one has to compute the correlation matrix  $(\mathbf{A}, \mathbf{A})$  and its inverse  $(\mathbf{A}, \mathbf{A})^{-1}$ . The inner product denotes average over the canonical distribution  $\exp(-\beta \mathcal{H}(\Gamma))$  with  $\beta \equiv 1/(k_B T)$ :

$$(\mathbf{a}, \mathbf{b}) \equiv \langle \mathbf{a}^* \mathbf{b} \rangle = \frac{1}{Z} \int d\Gamma \mathbf{a}^*(\Gamma) \mathbf{b}(\Gamma) e^{-\beta \mathcal{H}(\Gamma)} \quad (\text{III.7})$$

where  $Z$  is the partition function  $Z \equiv \int d\Gamma \exp(-\beta \mathcal{H}(\Gamma))$ .

### A. The correlation matrix $(\mathbf{A}, \mathbf{A})$

The correlation matrix  $\mathbf{C} = (\mathbf{A}, \mathbf{A})$  is given by

$$(\mathbf{A}, \mathbf{A}) = \begin{pmatrix} (\Delta\mathbf{R}_\alpha, \Delta\mathbf{R}_\beta) & (\mathbf{P}_\alpha, \Delta\mathbf{R}_\beta) & (\delta\rho_{\mathbf{k}}^a, \Delta\mathbf{R}_\beta) & (\mathbf{J}_{\mathbf{k}}^a, \Delta\mathbf{R}_\beta) \\ (\Delta\mathbf{R}_\alpha, \mathbf{P}_\beta) & (\mathbf{P}_\alpha, \mathbf{P}_\beta) & (\delta\rho_{\mathbf{k}}^a, \mathbf{P}_\beta) & (\mathbf{J}_{\mathbf{k}}^a, \mathbf{P}_\beta) \\ (\Delta\mathbf{R}_\alpha, \delta\rho_{\mathbf{k}}^b) & (\mathbf{P}_\alpha, \delta\rho_{\mathbf{k}}^b) & (\delta\rho_{\mathbf{k}}^a, \delta\rho_{\mathbf{k}}^b) & (\mathbf{J}_{\mathbf{k}}^a, \delta\rho_{\mathbf{k}}^b) \\ (\Delta\mathbf{R}_\alpha, \mathbf{J}_{\mathbf{k}}^b) & (\mathbf{P}_\alpha, \mathbf{J}_{\mathbf{k}}^b) & (\delta\rho_{\mathbf{k}}^a, \mathbf{J}_{\mathbf{k}}^b) & (\mathbf{J}_{\mathbf{k}}^a, \mathbf{J}_{\mathbf{k}}^b) \end{pmatrix} \quad (\text{III.8})$$

We first identify the vanishing elements. The following elements vanish:

$$\begin{aligned} (\Delta\mathbf{R}_\alpha, \mathbf{P}_\beta) &= 0, & (\Delta\mathbf{R}_\alpha, \mathbf{J}_{\mathbf{k}}^b) &= 0 \\ (\mathbf{P}_\alpha, \Delta\mathbf{R}_\beta) &= 0, & (\mathbf{P}_\alpha, \delta\rho_{\mathbf{k}}^b) &= 0, & (\mathbf{P}_\alpha, \mathbf{J}_{\mathbf{k}}^b) &= 0, \\ (\delta\rho_{\mathbf{k}}^a, \mathbf{P}_\beta) &= 0, & (\delta\rho_{\mathbf{k}}^a, \mathbf{J}_{\mathbf{k}}^b) &= 0, \\ (\mathbf{J}_{\mathbf{k}}^a, \Delta\mathbf{R}_\beta) &= 0, & (\mathbf{J}_{\mathbf{k}}^a, \mathbf{P}_\beta) &= 0, & (\mathbf{J}_{\mathbf{k}}^a, \delta\rho_{\mathbf{k}}^b) &= 0 \end{aligned} \quad (\text{III.9})$$

They vanish since the momentum integrations

$$\int d\mathbf{p}^{nN} \mathbf{p}_i^a \exp(-\beta \sum_i \sum_a \frac{\mathbf{p}_i^a \cdot \mathbf{p}_i^a}{2m_a}) = 0, \quad \int d\mathbf{P}^{N_u} \mathbf{P}_\alpha \exp(-\beta \sum_\gamma \mathbf{P}_\gamma^2 / 2M_\gamma) = 0.$$

We now look at the nonvanishing elements. The momentum correlation of solute particles is easy to compute:

$$\begin{aligned} (\mathbf{P}_\alpha, \mathbf{P}_\beta) &= \frac{1}{Z_P} \int d\mathbf{P}^{N_u} \mathbf{P}_\alpha \mathbf{P}_\beta e^{-\beta \sum_\gamma \mathbf{P}_\gamma^2 / 2M_\gamma} \\ &= \frac{1}{Z_P} \int d\mathbf{P}^{N_u} \mathbf{P}_\alpha (-M_\beta k_B T) \frac{\partial}{\partial \mathbf{P}_\beta} e^{-\beta \sum_\gamma \mathbf{P}_\gamma^2 / 2M_\gamma} = k_B T M_\alpha \mathbf{1} \delta_{\alpha\beta} \end{aligned} \quad (\text{III.10})$$

where  $Z_P \equiv \int d\mathbf{P}^{N_u} \exp(-\beta \sum_\gamma \mathbf{P}_\gamma^2 / 2M_\gamma)$ , and  $\mathbf{1}$  is the unit  $(3 \times 3)$  matrix. The eq. (III.10) is nothing but the equipartition theorem.

Since the general current-current correlation function  $(\mathbf{J}_{\mathbf{k}}^a, \mathbf{J}_{\mathbf{k}}^b)$  will have non-vanishing correlation between the same Cartesian components only, it is sufficient to define the current-current correlation function as

$$J_{ab}(k) \equiv \frac{1}{N} \langle \mathbf{J}_{-\mathbf{k}}^a \cdot \mathbf{J}_{\mathbf{k}}^b \rangle \quad (\text{III.11})$$

Its calculation is somewhat involved:

$$J_{ab}(k) = \frac{1}{N} \sum_i \sum_j \frac{1}{m_a m_b} \langle \mathbf{p}_i^a \cdot \mathbf{p}_j^b e^{-i\mathbf{k} \cdot (\mathbf{r}_i^a - \mathbf{r}_j^b)} \rangle$$

$$\begin{aligned}
&= \frac{1}{N} \sum_i \sum_j \frac{1}{m_a m_b} \langle \mathbf{p}_i^a \cdot \mathbf{p}_j^b \rangle \langle e^{-i\mathbf{k} \cdot (\mathbf{r}_i^a - \mathbf{r}_j^b)} \rangle \\
&= \frac{1}{N} \sum_i \sum_j \langle \mathbf{v}_i^a \cdot \mathbf{v}_i^b \rangle \delta_{ij} \langle e^{-i\mathbf{k} \cdot (\mathbf{r}_i^a - \mathbf{r}_j^b)} \rangle \\
&= \frac{1}{N} \sum_i \langle \mathbf{v}_i^a \cdot \mathbf{v}_i^b \rangle \langle e^{-i\mathbf{k} \cdot (\mathbf{r}_i^a - \mathbf{r}_i^b)} \rangle
\end{aligned} \tag{III.12}$$

A general expression of this quantity is given in Eq. (7) in [18].

The remaining elements  $(\Delta\mathbf{R}_\alpha, \Delta\mathbf{R}_\beta)$ ,  $(\Delta\mathbf{R}_\alpha, \delta\rho_{\mathbf{k}}^b)$ ,  $(\delta\rho_{\mathbf{k}}^a, \Delta\mathbf{R}_\beta)$ , and  $(\delta\rho_{\mathbf{k}}^a, \delta\rho_{\mathbf{k}}^b)$  involve the spatial coordinates only. We consider them in order. In the present work we will not specify particular form for correlation of initial position of solute particles since here we are interested in laying out general structure of the dynamics. We first have the displacement correlation matrix for the solute particles

$$\mathbf{L}_{\alpha\beta} \equiv (\Delta\mathbf{R}_\alpha, \Delta\mathbf{R}_\beta) \tag{III.13}$$

where  $\mathbf{L}_{\alpha\beta}$  is a  $(3N_u \times 3N_u)$  matrix. We next consider  $(\Delta\mathbf{R}_\alpha, \delta\rho_{\mathbf{k}}^b)$ . First note that  $(\Delta\mathbf{R}_\alpha, \delta\rho_{\mathbf{k}}^b) = \langle \Delta\mathbf{R}_\alpha \rho_{\mathbf{k}}^b \rangle - (2\pi)^3 \rho_0^b \delta(\mathbf{k}) \langle \Delta\mathbf{R}_\alpha \rangle = \langle \Delta\mathbf{R}_\alpha \rho_{\mathbf{k}}^b \rangle$  since  $\langle \Delta\mathbf{R}_\alpha \rangle = 0$ . Therefore

$$\mathbf{B}_{\mathbf{k}}^{\alpha,b} \equiv (\Delta\mathbf{R}_\alpha, \delta\rho_{\mathbf{k}}^b) = \langle \Delta\mathbf{R}_\alpha \rho_{\mathbf{k}}^b \rangle \tag{III.14}$$

In Appendix A, we show that this quantity and its transposed one vanish;

$$\mathbf{B}_{\alpha,b}(\mathbf{k}) = \mathbf{B}_{a,\beta}(\mathbf{k}) = \mathbf{0} \tag{III.15}$$

Finally, we have the static structure factor of solvent molecular liquid defined as

$$\chi_{ab}(\mathbf{k}) \equiv \frac{1}{N} \langle \delta\rho_{\mathbf{k}}^a \delta\rho_{\mathbf{k}}^b \rangle = \frac{1}{N} \langle \delta\rho_{-\mathbf{k}}^a \delta\rho_{\mathbf{k}}^b \rangle \tag{III.16}$$

This can be calculated using the RISM theory.

Summing up the above results, we have the following block-diagonal matrix for  $(\mathbf{A}, \mathbf{A})$ .

$$(\mathbf{A}, \mathbf{A}) = \begin{pmatrix} \mathbf{L}_{\alpha\beta} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & k_B T M_\alpha \mathbf{1} \delta_{\alpha\beta} & \mathbf{0} & \mathbf{0} \\ \mathbf{0}^T & \mathbf{0}^T & N \chi_{ab}(\mathbf{k}) & \underline{\mathbf{0}} \\ \mathbf{0}^T & \mathbf{0}^T & \underline{\mathbf{0}}^T & N J_{ab}(k) \end{pmatrix} \tag{III.17}$$

where  $\mathbf{O}$  denotes the  $(3N_u \times 3N_u)$  zero matrix,  $\mathbf{0}$  the  $(3N_u \times n)$  zero matrix,  $\underline{\mathbf{0}}$  the  $(n \times n)$  zero matrix, and the superscript  $T$  the transpose matrix.

### B. Inverse of $(\mathbf{A}, \mathbf{A})$

Since the above correlation matrix is block-diagonal, it is trivial to obtain the inverse  $(\mathbf{A}, \mathbf{A})^{-1}$  as

$$(\mathbf{A}, \mathbf{A})^{-1} = \begin{pmatrix} (\mathbf{L}^{-1})_{\alpha\beta} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{k_B T M_\alpha} \mathbf{1} \delta_{\alpha\beta} & \mathbf{0} & \mathbf{0} \\ \mathbf{0}^T & \mathbf{0}^T & \frac{1}{N} \chi_{ab}^{-1}(k) & \underline{\mathbf{0}} \\ \mathbf{0}^T & \mathbf{0}^T & \mathbf{0} & \frac{1}{N} J_{ab}^{-1}(k) \end{pmatrix} \quad (\text{III.18})$$

Here the inverse matrices  $(\mathbf{L}^{-1})_{\alpha\beta}$ ,  $\chi_{ab}^{-1}(k)$  and  $J_{ab}^{-1}(k)$  are defined as

$$\sum_{\gamma=1}^{N_u} \mathbf{L}_{\alpha\gamma} (\mathbf{L}^{-1})_{\gamma\beta} = \mathbf{1} \delta_{\alpha\beta}, \quad \sum_{c=1}^n \chi_{ac}(k) \chi_{cb}^{-1}(k) = \delta_{ab}, \quad \sum_{c=1}^n J_{ac}(k) J_{cb}^{-1}(k) = \delta_{ab}.$$

### C. The frequency matrix $i\Omega$

Here we compute the frequency matrix  $i\Omega$  which is defined as

$$i\Omega_{\lambda\nu} \equiv \sum_{\lambda'} (\mathbf{A}_{\lambda'}, \dot{\mathbf{A}}_\lambda) [(\mathbf{A}, \mathbf{A})^{-1}]_{\lambda'\nu} \quad (\text{III.19})$$

We first look at the elements of the matrix  $(\mathbf{A}, \dot{\mathbf{A}})$ :

$$(\mathbf{A}, \dot{\mathbf{A}}) = \begin{pmatrix} (\Delta \mathbf{R}_\alpha, \Delta \dot{\mathbf{R}}_\beta) & (\mathbf{P}_\alpha, \Delta \dot{\mathbf{R}}_\beta) & (\delta \rho_{\mathbf{k}}^a, \Delta \dot{\mathbf{R}}_\beta) & (\mathbf{J}_{\mathbf{k}}^a, \Delta \dot{\mathbf{R}}_\beta) \\ (\Delta \mathbf{R}_\alpha, \dot{\mathbf{P}}_\beta) & (\mathbf{P}_\alpha, \dot{\mathbf{P}}_\beta) & (\delta \rho_{\mathbf{k}}^a, \dot{\mathbf{P}}_\beta) & (\mathbf{J}_{\mathbf{k}}^a, \dot{\mathbf{P}}_\beta) \\ (\Delta \mathbf{R}_\alpha, \dot{\rho}_{\mathbf{k}}^b) & (\mathbf{P}_\alpha, \dot{\rho}_{\mathbf{k}}^b) & (\delta \rho_{\mathbf{k}}^a, \dot{\rho}_{\mathbf{k}}^b) & (\mathbf{J}_{\mathbf{k}}^a, \dot{\rho}_{\mathbf{k}}^b) \\ (\Delta \mathbf{R}_\alpha, \dot{\mathbf{J}}_{\mathbf{k}}^b) & (\mathbf{P}_\alpha, \dot{\mathbf{J}}_{\mathbf{k}}^b) & (\delta \rho_{\mathbf{k}}^a, \dot{\mathbf{J}}_{\mathbf{k}}^b) & (\mathbf{J}_{\mathbf{k}}^a, \dot{\mathbf{J}}_{\mathbf{k}}^b) \end{pmatrix} \quad (\text{III.20})$$

First we obtain some elements of  $\dot{\mathbf{A}}$  using the Liouville operator (III.2).

$$\begin{aligned} \Delta \dot{\mathbf{R}}_\alpha &= i\mathcal{L} \Delta \mathbf{R}_\alpha = \frac{\mathbf{P}_\alpha}{M_\alpha} \\ \dot{\mathbf{P}}_\alpha &= i\mathcal{L} \mathbf{P}_\alpha = \mathbf{F}_\alpha \\ \dot{\rho}_{\mathbf{k}}^a &= i\mathcal{L} \rho_{\mathbf{k}}^a = i\mathbf{k} \cdot \mathbf{J}_{\mathbf{k}}^a \\ \dot{\mathbf{J}}_{\mathbf{k}}^a &= \sum_i \frac{1}{m_a} \left( \dot{\mathbf{p}}_i^a + \mathbf{p}_i^a i\mathbf{k} \cdot \frac{\mathbf{p}_i^a}{m_a} \right) e^{i\mathbf{k} \cdot \mathbf{r}_i^a} \end{aligned} \quad (\text{III.21})$$

where  $\mathbf{F}_\alpha$  is the total force exerted on the  $\alpha$ th solute particle by the solvent as well as by other solute particles. Actually when we compute the elements involving  $\dot{\mathbf{P}}$  or  $\dot{\mathbf{J}}_{\mathbf{k}}^a$ , it is more convenient to use the integration by parts. It is useful to remember that whereas  $\Delta \dot{\mathbf{R}}$  and  $\dot{\rho}_{\mathbf{k}}^a$  involve single momentum ( $\mathbf{P}$  or  $\mathbf{p}_i$ ),  $\dot{\mathbf{P}}$  and  $\dot{\mathbf{J}}_{\mathbf{k}}^a$  involve zero (since  $\mathbf{p}_i^a$  is the force acting on the  $a$ th atom of the  $i$ th molecule, which

only involves the positions of solute particles and solvent molecules), or two momentums (two  $\mathbf{p}_i$ ). Using this fact, we can easily identify the vanishing elements:

$$\begin{aligned} (\Delta \mathbf{R}_\alpha, \Delta \dot{\mathbf{R}}_\beta) &= 0, \quad (\Delta \mathbf{R}_\alpha, \dot{\rho}_{\mathbf{k}}^b) = 0, \quad (\Delta \mathbf{R}_\alpha, \dot{\mathbf{J}}_{\mathbf{k}}^b) = 0 \\ (\mathbf{P}_\alpha, \dot{\mathbf{P}}_\beta) &= 0, \quad (\mathbf{P}_\alpha, \dot{\rho}_{\mathbf{k}}^a) = 0, \quad (\mathbf{P}_\alpha, \dot{\mathbf{J}}_{\mathbf{k}}^b) = 0, \\ (\delta \rho_{\mathbf{k}}^a, \Delta \dot{\mathbf{R}}_\beta) &= 0, \quad (\delta \rho_{\mathbf{k}}^a, \dot{\mathbf{P}}_\beta) = 0, \quad (\delta \rho_{\mathbf{k}}^a, \dot{\rho}_{\mathbf{k}}^b) = 0, \\ (\mathbf{J}_{\mathbf{k}}^a, \Delta \dot{\mathbf{R}}_\beta) &= 0, \quad (\mathbf{J}_{\mathbf{k}}^a, \dot{\mathbf{P}}_\beta) = 0, \quad (\mathbf{J}_{\mathbf{k}}^a, \dot{\mathbf{J}}_{\mathbf{k}}^b) = 0 \end{aligned} \quad (\text{III.22})$$

The nonvanishing elements are

$$\begin{aligned} (\Delta \mathbf{R}_\alpha, \dot{\mathbf{P}}_\beta) &= -\frac{1}{M}(\mathbf{P}_\alpha, \mathbf{P}_\beta) = -k_B T \mathbf{1} \delta_{\alpha\beta} \\ (\mathbf{P}_\alpha, \Delta \dot{\mathbf{R}}_\beta) &= \frac{1}{M}(\mathbf{P}_\alpha, \mathbf{P}_\beta) = k_B T \mathbf{1} \delta_{\alpha\beta} \\ (\dot{\rho}_{\mathbf{k}}^a, \dot{\mathbf{J}}_{\mathbf{k}}^b) &= (\dot{\rho}_{\mathbf{k}}^a, \mathbf{J}_{\mathbf{k}}^b) = i \mathbf{k} \cdot (\mathbf{J}_{\mathbf{k}}^a, \dot{\mathbf{J}}_{\mathbf{k}}^b) = i N \mathbf{k} J_{ab}(k) \\ (\mathbf{J}_{\mathbf{k}}^a, \dot{\rho}_{\mathbf{k}}^b) &= (\mathbf{J}_{\mathbf{k}}^a, \mathbf{J}_{\mathbf{k}}^b) \cdot i \mathbf{k} = i N \mathbf{k} J_{ab}(k) \end{aligned} \quad (\text{III.23})$$

Taking all these into account, we obtain

$$i\Omega = \begin{pmatrix} \mathbf{0} & k_B T \mathbf{1} \delta_{\alpha\beta} & \mathbf{0} & \mathbf{0} \\ -k_B T \mathbf{1} \delta_{\alpha\beta} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0}^T & \mathbf{0}^T & \underline{\mathbf{0}} & i N \mathbf{k} J_{ab}(k) \\ \mathbf{0}^T & \mathbf{0}^T & i N \mathbf{k} J_{ab}(k) & \underline{\mathbf{0}} \end{pmatrix} \cdot (\mathbf{A}, \mathbf{A})^{-1} \quad (\text{III.24})$$

Using the inverse correlation matrix (III.18), we compute  $i\Omega$  as

$$i\Omega = \begin{pmatrix} \mathbf{0} & \frac{1}{M_\alpha} \mathbf{1} \delta_{\alpha\beta} & \mathbf{0} & \mathbf{0} \\ -k_B T (\mathbf{L}^{-1})_{\alpha\beta} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0}^T & \mathbf{0}^T & \underline{\mathbf{0}} & i \mathbf{k} \delta_{ab} \\ \mathbf{0}^T & \mathbf{0}^T & i \mathbf{k} \sum_{c=1}^n J_{ac}(k) \chi_{cb}^{-1}(k) & \underline{\mathbf{0}} \end{pmatrix} \quad (\text{III.25})$$

#### D. The reversible part

From (II.5), the reversible part of the Langevin equation is given by  $i\Omega \cdot \mathbf{A}(t)$ . Using (III.25), we obtain

$$i\Omega \cdot \mathbf{A}(t) = \begin{pmatrix} \mathbf{P}_\alpha(t)/M_\alpha \\ -k_B T \sum_\beta (\mathbf{L}^{-1})_{\alpha\beta} \cdot \Delta \mathbf{R}_\beta(t) \\ i \mathbf{k} \cdot \mathbf{J}_{\mathbf{k}}^a(t) \\ i \mathbf{k} \sum_{b,c} J_{ac}(k) \chi_{cb}^{-1}(k) \delta \rho_{\mathbf{k}}^b(t) \end{pmatrix} \quad (\text{III.26})$$

### E. The fluctuating force

The fluctuating force at  $t = 0$  from (II.6) is given by

$$\mathbf{f} = (1 - \mathcal{P})\dot{\mathbf{A}} = \dot{\mathbf{A}} - i\boldsymbol{\Omega} \cdot \mathbf{A} \quad (\text{III.27})$$

where we used  $\mathcal{P}\dot{\mathbf{A}} = (\mathbf{A}, \dot{\mathbf{A}}) \cdot (\mathbf{A}, \mathbf{A})^{-1} \mathbf{A} = i\boldsymbol{\Omega} \cdot \mathbf{A}$ . We first obtain

$$\dot{\mathbf{A}} = \begin{pmatrix} \Delta\dot{\mathbf{R}}_\alpha \\ \dot{\mathbf{P}}_\alpha \\ \dot{\rho}_{\mathbf{k}}^a \\ \dot{\mathbf{J}}_{\mathbf{k}}^a \end{pmatrix} = \begin{pmatrix} \mathbf{P}_\alpha/M_\alpha \\ \mathbf{F}_\alpha \\ i\mathbf{k} \cdot \mathbf{J}_{\mathbf{k}}^a \\ \dot{\mathbf{J}}_{\mathbf{k}}^a \end{pmatrix} \quad (\text{III.28})$$

and

$$i\boldsymbol{\Omega} \cdot \mathbf{A} = \begin{pmatrix} \mathbf{P}_\alpha/M_\alpha \\ -k_B T \sum_\beta (\mathbf{L}^{-1})_{\alpha\beta} \cdot \Delta\mathbf{R}_\beta \\ i\mathbf{k} \cdot \mathbf{J}_{\mathbf{k}}^a \\ i\mathbf{k} \sum_{b,c} J_{ac}(k) \chi_{cb}^{-1}(k) \delta\rho_{\mathbf{k}}^b \end{pmatrix} \quad (\text{III.29})$$

which is obtained by setting  $t = 0$  in (III.26). Using the above two results, we obtain for the fluctuating force as

$$\mathbf{f} = \begin{pmatrix} \mathbf{0} \\ \mathbf{W}_\alpha \\ 0 \\ \Xi_{\mathbf{k}}^a \end{pmatrix}, \quad \mathbf{f}(t) = e^{it(1-\mathcal{P})\mathcal{L}} \begin{pmatrix} \mathbf{0} \\ \mathbf{W}_\alpha \\ 0 \\ \Xi_{\mathbf{k}}^a \end{pmatrix} \quad (\text{III.30})$$

where

$$\mathbf{W}_\alpha \equiv \mathbf{F}_\alpha + k_B T \sum_\beta (\mathbf{L}^{-1})_{\alpha\beta} \cdot \Delta\mathbf{R}_\beta, \quad \Xi_{\mathbf{k}}^a \equiv \dot{\mathbf{J}}_{\mathbf{k}}^a - i\mathbf{k} \sum_{b,c} J_{ac}(k) \chi_{cb}^{-1}(k) \delta\rho_{\mathbf{k}}^b \quad (\text{III.31})$$

### F. The memory matrix

The memory function matrix  $\mathbf{K}(t)$  is calculated as

$$\begin{aligned} \mathbf{K}(t) &\equiv (\mathbf{f}, \mathbf{f}(t))[(\mathbf{A}, \mathbf{A})^{-1}] \\ &= \begin{pmatrix} \mathbf{O} & \mathbf{O} & \mathbf{0} \\ \mathbf{O} & \frac{1}{k_B T M_\alpha} (\mathbf{W}_\alpha, \mathbf{W}_\beta(t)) & \mathbf{0} & \frac{1}{N} \sum_b J_{ab}^{-1}(k) (\Xi_{\mathbf{k}}^b, \mathbf{W}_\beta(t)) \\ \mathbf{0}^T & \mathbf{0}^T & \mathbf{0} \\ \mathbf{0}^T & \frac{1}{M k_B T} (\mathbf{W}_\alpha, \Xi_{\mathbf{k}}^b(t)) & \mathbf{0} & \frac{1}{N} \sum_c J_{ac}^{-1}(k) (\Xi_{\mathbf{k}}^c, \Xi_{\mathbf{k}}^b(t)) \end{pmatrix} \end{aligned} \quad (\text{III.32})$$

where  $\mathbf{W}(t) \equiv \exp(it(1 - \mathcal{P})\mathcal{L})\mathbf{W}$  and  $\Xi_{\mathbf{k}}^a(t) \equiv \exp(it(1 - \mathcal{P})\mathcal{L})\Xi_{\mathbf{k}}^a$ .

In (III.32), the two terms exhibit explicit  $N$ -dependence. In the thermodynamic limit in which  $N$  is taken to be infinite while the number of solute particles  $N_u$  remains finite, the term  $\frac{1}{N} \sum_b J_{ab}^{-1}(k)(\Xi_{\mathbf{k}}^b, \mathbf{W}(t))$  will vanish since the ensemble average  $(\Xi_{\mathbf{k}}, \mathbf{W}(t))$  will remain finite. The other term  $\frac{1}{N} \sum_c J_{ac}^{-1}(k)(\Xi_{\mathbf{k}}^c, \Xi_{\mathbf{k}}^b(t))$  will not vanish since the ensemble average  $(\Xi_{\mathbf{k}}^c, \Xi_{\mathbf{k}}^b(t))$  is proportional to  $N$ . Therefore only the latter term survives in the thermodynamic limit.

In Appendix B, we show that

$$(\mathbf{W}_{\alpha}, \Xi_{\mathbf{k}}^b(t)) = 0 \quad (\text{III.33})$$

Therefore the final expression of the memory matrix is given by

$$\mathbf{K}(t) = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{k_B T M_{\alpha}}(\mathbf{W}_{\alpha}, \mathbf{W}_{\beta}(t)) & \mathbf{0} \\ \mathbf{0}^T & \mathbf{0}^T & 0 \\ \mathbf{0}^T & \mathbf{0}^T & \frac{1}{N} \sum_c J_{ac}^{-1}(k)(\Xi_{\mathbf{k}}^c, \Xi_{\mathbf{k}}^b(t)) \end{pmatrix} \quad (\text{III.34})$$

#### G. The explicit form of the exact dynamic equations

With the explicit results of the previous sections, we here write down an explicit form for the time evolution equation (II.5)

$$\begin{aligned} \frac{d\Delta\mathbf{R}_{\alpha}(t)}{dt} &= \frac{\mathbf{P}_{\alpha}(t)}{M_{\alpha}}, \\ \frac{d\mathbf{P}_{\alpha}(t)}{dt} &= -k_B T \sum_{\beta} (\mathbf{L}^{-1})_{\alpha\beta} \cdot \Delta\mathbf{R}_{\beta}(t) - \int_0^t ds \sum_{\beta} \boldsymbol{\Gamma}_{\alpha\beta}(t-s) \cdot \frac{\mathbf{P}_{\beta}(s)}{M_{\beta}} + \mathbf{W}_{\alpha}(t), \\ \frac{d\delta\rho_{\mathbf{k}}^a(t)}{dt} &= i\mathbf{k} \cdot \mathbf{J}_{\mathbf{k}}^a(t), \\ \frac{d\mathbf{J}_{\mathbf{k}}^a(t)}{dt} &= i\mathbf{k} \sum_{b,c} J_{ac}(k) \chi_{cb}^{-1}(k) \delta\rho_{\mathbf{k}}^b(t) - \frac{1}{N} \sum_{b,c} J_{ac}^{-1}(k) \int_0^t ds \mathbf{M}_{\mathbf{k}}^{bc}(t-s) \cdot \mathbf{J}_{\mathbf{k}}^b(s) + \Xi_{\mathbf{k}}^a(t) \end{aligned} \quad (\text{III.35})$$

In the above set of dynamic equations for the solute and solvent molecules, the random forces take the following forms

$$\mathbf{W}_{\alpha}(t) = e^{it(1-\mathcal{P})\mathcal{L}} \left( \mathbf{F}_{\alpha} + k_B T \sum_{\beta} (\mathbf{L}^{-1})_{\alpha\beta} \cdot \Delta\mathbf{R}_{\beta} \right),$$

$$\Xi_{\mathbf{k}}^a(t) = e^{it(1-\mathcal{P})\mathcal{L}} \left( \dot{\mathbf{J}}_{\mathbf{k}}^a - i\mathbf{k} \sum_{b,c} J_{ac}(k) \chi_{cb}^{-1}(k) \delta\rho_{\mathbf{k}}^b \right) \quad (\text{III.36})$$

The memory functions in (III.35) are given by the time correlations of the random forces;

$$\Gamma_{\alpha\beta}(t) = \frac{1}{k_B T} \langle \mathbf{W}_{\alpha}(t) \mathbf{W}_{\beta}(0) \rangle, \quad \mathbf{M}_{\mathbf{k}}^{bc}(t) = \langle \Xi_{\mathbf{k}}^b(t) \Xi_{-\mathbf{k}}^c(0) \rangle. \quad (\text{III.37})$$

## IV. DISCUSSIONS

### A. Solvent dynamics

When the fluid is far from protein, or in bulk, where perturbation from protein vanishes, the last two expressions concerning solvent in Eq. (III.35) reduce to the equations for pure-water dynamics, Eqs. (24) and (25), derived by one of the authors [19], except for an approximation made in the factor  $Jab$  [20]. The equation is further simplified to produce the site-site Smoluchowski-Vlasov (SSSV) equation if one makes the memory function local in time as well as in space. The equation can be analytically solved by means of the Laplace transform to produce the van Hove or space-time correlation function of water, with the *input* of the site-site pair correlation functions of the solvent obtained from the RISM theory. The theory has been successfully applied to a variety of solvent relaxation processes induced by an abrupt change in the electronic structure of a solute molecule, or solvation dynamics, which can be probed by the dynamic Stokes-shift [21]. So, the two equations concerning solvent in (III.35) can be regarded as a generalization of the previous theories developed for pure water to that subject to the field exerted from protein atoms. There are several remarks to be made with respect to the generalization. Firstly, the translational invariance of the system is no longer valid. Therefore, the equations should be solved in three-dimensional Cartesian-space. Secondly, the factor  $\chi_{ab}(\mathbf{r}, \mathbf{r}') = N^{-1} \langle \delta\rho^a(\mathbf{r}) \delta\rho^b(\mathbf{r}') \rangle$  appearing in the equation is a two body density correlation function, but subject to the "external force" due to protein. Such a theory for obtaining the function is under development, but it is too primitive at the moment to be applied to the problem we are facing. Therefore, we may adopt the superposition approximation  $\langle \delta\rho^a(\mathbf{r}) \delta\rho^b(\mathbf{r}') \rangle = \langle \delta\rho^a(\mathbf{r}) \rangle \langle \delta\rho^b(\mathbf{r}') \rangle$  to this case. Then,  $\langle \delta\rho^a(\mathbf{r}) \rangle$  can be readily evaluated from the 3D-RISM theory.

A number of possible applications of the dynamic equations for the solvent are conceivable. An interesting example is the current-current correlation function  $\langle \mathbf{J}^a(\mathbf{r}, 0) \mathbf{J}^b(\mathbf{r}', t) \rangle$  of water and ions in a molecular channel, which is concerned with many observables including the permeability of water and ions across the cell membrane [22]. The equation for the correlation function can be readily obtained by coupling the two equations for solvent with the aid of the 3D-RISM/RISM theory.

## B. Solute dynamics

The first two equations in Eq. (III.35) concerning solute dynamics are combined together to result in

$$M_\alpha \frac{d^2 \Delta \mathbf{R}_\alpha(t)}{dt^2} + k_B T \sum_\beta (\mathbf{L}^{-1})_{\alpha\beta} \cdot \Delta \mathbf{R}_\beta(t) + \int_0^t ds \sum_\beta \Gamma_{\alpha\beta}(t-s) \cdot \frac{d\Delta \mathbf{R}_\beta(s)}{ds} = \mathbf{W}_\alpha(t) \quad (\text{IV.1})$$

The equation is regarded as a generalization of the equation for a coupled set of Langevin-oscillators, first examined by Wang and Uhlenbeck [23], to a realistic model of protein in water. Wang and Uhlenbeck proposed a model in which a coupled set of oscillators consisting of spherical beads is immersed in a viscous liquid, and applied the Langevin theory to the oscillators. Later on, Lamm and Szabo [24] performed a normal mode analysis on the Wang-Uhlenbeck oscillators, assuming a phenomenological friction term. Kottlam and Case [25], and Ansari [26] applied the Langevin mode method of Lamm and Szabo to proteins. The same method was also applied to the dynamics of DNA [27] and RNA [28] in solvents. A review on the normal mode analysis in general (including the Langevin mode analysis) in the dynamics of biomolecules is presented in [29].

There are several comments to be made on the new equation (IV.1). Firstly, the equation does not include a term related to the force which originates from the first derivative of the free energy surface with respect to the position. The force acting on an atom of protein comes from the three contributions, one which is proportional to the displacement of the atom from its equilibrium position (the second term in the left hand side in (IV.1)), and the friction term proportional to the velocities of the atoms, and that due to the random force (the term in the right hand side in (IV.1)). The physical origin as to why the equation does not include the first derivative of the free energy lies in our treatment based on the generalized Langevin theory. The whole idea of the generalized Langevin theory is to project all the degrees of freedom in the phase space onto few dynamic variables under concern. The projection is carried out using a projection operator, defined by Eqs. (II.3) and (II.4), in terms of an ensemble average of two variables which are fluctuating around an equilibrium in the phase space. Obviously by definition, the ensemble average of the displacement of atoms in protein should be zero in equilibrium.

Such a force as the first derivative of the free energy, which may cause the complete shift of the equilibrium, is not included in the treatment. The situation is somewhat analogous to the case of a harmonic oscillator, in which an oscillator swings back and forth around a minimum of the harmonic potential. Only force acts on the system is the restoring force proportional to the displacement from the potential minimum. In our case, too, only force acting on the protein atoms is the one which

restores atom positions from fluctuating to equilibrium ones. However, there is an essential difference in physics between the two systems. The equilibrium position of a harmonic oscillator is the minimum of mechanical potential energy, while that of protein in water is the minimum in the thermodynamic potential or the free energy, which is concerned not only with energy but also with the entropy both of protein and of water. So, in the case of protein in water, the stochastic character of the dynamics is attributed not only to the random force term, but also to the conformational fluctuation of protein around its equilibrium state, induced by solvent, while the stochastic character is resulted just from the random force term in the case of the coupled harmonic oscillators treated by Wang and Unlenbeck.

The argument above suggests interesting physics implied in Eq. (IV.1), and its application to biological functions. If one ignores the friction and random force terms in Eq. (IV.1), one gets

$$M_\alpha \frac{d^2 \Delta \mathbf{R}_\alpha(t)}{dt^2} = -k_B T \sum_\beta (\mathbf{L}^{-1})_{\alpha\beta} \cdot \Delta \mathbf{R}_\beta(t) \quad (\text{IV.2})$$

This equation can be viewed as a coupled set of "harmonic oscillators", whose "Hessian" matrix is given by  $k_B T (\mathbf{L}^{-1})_{\alpha\beta}$ . Considering Eq. (III.13), the "Hessian" matrix is related to the variance-covariance matrix of the positional fluctuation by

$$k_B T \mathbf{L}^{-1} \equiv k_B T \langle \Delta \mathbf{R} \Delta \mathbf{R} \rangle^{-1} \quad (\text{IV.3})$$

The observation strongly suggests that the dynamics described by Eq. (IV.2) is that of fluctuation around a minimum of the free energy surface consisting not only of the interactions among atoms in the protein, but of the solvation free energy. In this respect, the configuration corresponding to the free energy minimum is not just one but an ensemble of distinguishable configurations concerning protein and solvent, which can be converted among each other due to the thermal noise. The free energy surface can be given by

$$F(\{\Delta \mathbf{R}\}) = U(\{\Delta \mathbf{R}\}) + \Delta \mu(\{\Delta \mathbf{R}\}) \quad (\text{IV.4})$$

where  $U(\{\Delta \mathbf{R}\})$  is the interaction potential energy among atoms in a protein, and  $\Delta \mu(\{\Delta \mathbf{R}\})$  is the solvation free energy of protein whose conformation is  $\{\mathbf{R}\}$  [30].

The above consideration further suggests a method to evaluate the variance-covariance matrix, which characterizes structural fluctuation of protein, based on the 3D-RISM theory. The variance-covariance matrix is closely related to the Hessian matrix, Eq. (IV.3), and the Hessian matrix is the second derivative of the free energy surface, namely,

$$k_B T (\mathbf{L}^{-1})_{\alpha\beta} = \frac{\partial^2 F(\{\Delta \mathbf{R}\})}{\partial \Delta \mathbf{R}_\alpha \partial \Delta \mathbf{R}_\beta} \quad (\text{IV.5})$$

Since the free energy  $F(\{\Delta\mathbf{R}\})$  can be obtained by solving the 3D-RISM/RISM equation, Eq.(IV.5) provides a way to evaluate the variance-covariance matrix.

The variance-covariance matrix is by itself quite informative for characterizing the structural fluctuation of protein around its native state in atomic detail. As an example, let us consider a hinge-bending motion of protein. The variance-covariance matrix should have a structure in which a block of elements  $\langle \Delta\mathbf{R}_\alpha \Delta\mathbf{R}_\beta \rangle$  for atom pairs,  $\alpha, \beta$ , belonging to the two sides of the hinge-axis, have the negative sign because the direction of the displacements  $\Delta\mathbf{R}_\alpha$  and  $\Delta\mathbf{R}_\beta$  is opposite.

Usefulness of the variance-covariance matrix is not limited to characterization of the structural fluctuation around an equilibrium state. The Eq. (IV.5) implies that the free energy of protein at an equilibrium conformation takes the form

$$F(\{\Delta\mathbf{R}\}) = \frac{1}{2}k_B T \sum_{\alpha, \beta} \Delta\mathbf{R}_\alpha \cdot (\mathbf{L}^{-1})_{\alpha\beta} \cdot \Delta\mathbf{R}_\beta. \quad (\text{IV.6})$$

In the presence of a small perturbation due to, say, ligand binding, the above free energy can be changed due to the perturbation as

$$F(\{\Delta\mathbf{R}\}) = \frac{1}{2}k_B T \sum_{\alpha, \beta} \Delta\mathbf{R}_\alpha \cdot (\mathbf{L}^{-1})_{\alpha\beta} \cdot \Delta\mathbf{R}_\beta - \sum_\alpha \Delta\mathbf{R}_\alpha \cdot \mathbf{f}_\alpha \quad (\text{IV.7})$$

where  $\mathbf{f}_\alpha$  is the force acting on the  $\alpha$ th protein atom due to the perturbation. Then, the conformational change due to the perturbation can be determined by the variational principle

$$\frac{\partial F}{\partial \Delta\mathbf{R}_\alpha} = 0. \quad (\text{IV.8})$$

With Eq. (IV.7), Eq. (IV.8) gives

$$\langle \Delta\mathbf{R}_\alpha \rangle_1 = (k_B T)^{-1} \sum_\beta \langle \Delta\mathbf{R}_\alpha \Delta\mathbf{R}_\beta \rangle_0 \cdot \mathbf{f}_\beta \quad (\text{IV.9})$$

where the subscript 1(0) denotes the presence (absence) of the perturbation. Therefore, Eq. (IV.5) combined with Eq.(IV.9) provides a theoretical basis for analyzing the conformational relaxation of protein in water due to a perturbation such as ligand binding. Th Eq.(IV.9) is first derived by Ikeguchi et. al. [7] based on the linear response theory.

The equation (IV.1) is also a generalized equation which provides molecular basis for the phenomenological Rouse-Zimm model of the polymer dynamics [31], with a proper account of the variance-covariance matrix, the diagonal terms of which correspond to the mean square displacement of each atom in equilibrium states. This suggests that the theory can be applied not only to the native conformation of protein but also to characterizing the denatured or random-coil state. However, the

application requires special care of the ensemble average to evaluate the variance-covariance matrix, since the average, by definition, should be taken over virtually an infinite number of conformations randomly appearing in the solution. Nevertheless, a practical method to evaluate the variance-covariance matrix for the random-coil state of protein can be suggested based on the 3D-RISM/RISM theory as follows. First, produce some small number of conformations for protein in water by means of a generalized ensemble technique such as the replica-exchange algorithm. Second, evaluate the second derivative of the free energy surface of each conformation based on Eq. (IV.5), and take the average of the results over the conformations, which will give rise to the variance-covariance matrix for the sampled conformational space. Third, add more conformations to the sample to take the average. Repeat the procedure until the convergence is attained. Our implication is that the convergence will be attained rather quickly, because the variance-covariance matrix for each conformation, obtained from Eq. (IV.5), is already an average over a large number of conformations in the free energy surface. The converged variance-covariance matrix can be compared with observable quantities which characterize a random coil state of protein, such as the gyration radius and the distribution of end-to-end distance.

## V. CONCLUDING REMARKS

In the present work, we have proposed a new theory of dynamics based on the generalized Langevin theory, which can be applied to structural fluctuation around a native state of protein in water. The displacement vector of atom positions and their conjugated momentum, are chosen for dynamic variables for protein, while the density fields of atoms and their momentum fields are chosen for water. Projection of other degrees of freedom onto those dynamic variables using the standard projection operator method produced essentially two equations which describe the time evolution of fluctuation concerning the density field of solvent and the conformation of protein around an equilibrium state, which are correlated each other.

The equation of motion for protein atoms in water is *formally* akin to that of the Langevin equation for coupled harmonic oscillators in the continuum solvent, examined by Wang and Uhlenbeck long time ago. However, there exists a substantial and important difference. Unlike the coupled set of harmonic oscillators, the "Hessian" included in the term corresponding to the Hookian-like restoring force in the new equation is identified as a variance-covariance matrix of the displacement vector, which is nothing but the second moment of the structural fluctuation. Since the fluctuation is taking place around the thermodynamic equilibrium, not just around a minimum of a (mechanical) harmonic

potential, the "Hessian" matrix should be related to the second derivative of the free energy surface of protein, which of course includes the influence of solvent.

All those findings suggest that we are now at the position where we can explore the conformational fluctuation around a native state of protein, correlated with the relaxation of water density, since a method to evaluate the free energy surface and its first derivative of protein in water has been well established already based on the "3D-RISM/RISM" theory. It is not difficult to calculate the second derivative of the free energy surface from the first derivatives.

The finding further suggests even more practical applications related to the drug design. The 3D-RISM/RISM theory has been successfully applied to a variety processes of molecular recognition in protein, including drug binding. However, so far the application has been limited to a fixed conformation of protein, which of course cannot take into account the effects of conformational fluctuation, such as the *induced fitting*. With the aid of the linear response theory, the new formulation provides a foundation to evaluate the effect of conformational fluctuation in the process of molecular recognition.

### Appendix A: Calculation of $\mathbf{B}_{\alpha,b}(\mathbf{k})$

We show that the correlation  $\mathbf{B}_{\alpha,b}(\mathbf{k}) \equiv \langle \Delta \mathbf{R}_\alpha \delta \rho_{\mathbf{k}}^b \rangle$  vanishes in the homogeneous system. An explicit expression for  $\mathbf{B}_{\alpha,b}(\mathbf{k})$  is given by

$$\mathbf{B}_{\alpha,b}(\mathbf{k}) \equiv \langle \Delta \mathbf{R}_\alpha \delta \rho_{\mathbf{k}}^b \rangle = \frac{1}{Z_c} \int d\Gamma_c \Delta \mathbf{R}_\alpha \sum_l e^{i\mathbf{k} \cdot \mathbf{r}_l^b} e^{-\beta \mathcal{U}} \quad (\text{A.1})$$

where  $\Gamma_c$  denotes collection of the position variables only, and  $Z_c$  is the corresponding partition function defined as  $Z_c \equiv \int d\Gamma_c \exp(-\beta \mathcal{U})$  with  $\mathcal{U}$  being the total potential energy of the system. Shifting integration variables in (A.1) as  $\mathbf{r}_l^b \rightarrow \mathbf{r}_l^b + \Delta \mathbf{R}_\alpha$  and  $\Delta \mathbf{R}_\beta \rightarrow \Delta \mathbf{R}_\beta + \Delta \mathbf{R}_\alpha$  ( $\beta \neq \alpha$ ), one can separate the  $\Delta \mathbf{R}_\alpha$ -integration as follows;

$$\langle \Delta \mathbf{R}_\alpha \delta \rho_{\mathbf{k}}^b \rangle = \frac{\int d\Delta \mathbf{R}_\alpha \Delta \mathbf{R}_\alpha e^{i\mathbf{k} \cdot \Delta \mathbf{R}_\alpha}}{\int d\Delta \mathbf{R}_\alpha} \cdot \frac{\int d\mathbf{R}^{N_u-1} \int d\mathbf{r}^{nN} \sum_l e^{i\mathbf{k} \cdot \mathbf{r}_l^b} e^{-\beta \mathcal{U}'}}{\int d\Delta \mathbf{R}^{N_u-1} \int d\mathbf{r}^{nN} e^{-\beta \mathcal{U}'}} \quad (\text{A.2})$$

where the potential energy  $\mathcal{U}'$  is obtained from  $\mathcal{U}$  with the shift of the variables, and does not involve  $\mathbf{R}_\alpha$ . We consider the first integration factor in (A.2). For simplicity we suppress the index  $\alpha$ . Its  $X$ -component is given by

$$\begin{aligned} \frac{\int d\mathbf{R} X e^{i\mathbf{k} \cdot \mathbf{R}}}{\int d\mathbf{R}} &= \frac{1}{L^3} \left( \int_{-L/2}^{L/2} X e^{ik_1 X} dX \right) \left( \int_{-L/2}^{L/2} e^{ik_2 Y} dY \right) \left( \int_{-L/2}^{L/2} e^{ik_3 Z} dZ \right) \\ &= \frac{1}{L^3} \cdot \frac{2}{k_1} \left[ \frac{1}{k_1} \sin\left(\frac{k_1 L}{2}\right) - \frac{L}{2} \cos\left(\frac{k_1 L}{2}\right) \right] \cdot \frac{2}{k_2} \sin\left(\frac{k_2 L}{2}\right) \cdot \frac{2}{k_3} \sin\left(\frac{k_3 L}{2}\right) \end{aligned} \quad (\text{A.3})$$

Note that this integral vanishes in the thermodynamic limit  $L \rightarrow \infty$  due to the oscillating term  $e^{i\mathbf{k} \cdot \mathbf{R}}$ . The second integration factor in (A.2) remains finite. Therefore we obtain  $\langle X \rho_{\mathbf{k}} \rangle = 0$  in the thermodynamic limit. Since this will hold for the other components, we conclude that

$$\mathbf{B}_{\alpha,b}(\mathbf{k}) = 0 \quad (\text{A.4})$$

### Appendix B: Calculation of $(\mathbf{W}_\alpha, \Xi_{\mathbf{k}}^b(t))$

Here we show that the memory matrix  $(\mathbf{W}_\alpha, \Xi_{\mathbf{k}}^b(t))$  vanishes. We consider

$$\langle \mathbf{W}_\alpha, \Xi_{\mathbf{k}}^b(t) \rangle = \langle \mathbf{W}_\alpha e^{it\mathcal{Q}\mathcal{L}} \Xi_{\mathbf{k}}^b \rangle = \langle \mathbf{W}_\alpha \left( 1 + t\mathcal{Q}i\mathcal{L} + \frac{t^2}{2!} \mathcal{Q}i\mathcal{L}\mathcal{Q}i\mathcal{L} + \dots \right) \Xi_{\mathbf{k}}^b \rangle \quad (\text{B.1})$$

where  $\mathcal{Q} = 1 - \mathcal{P}$ . For simplicity of notation, we write  $\mathbf{W}_\alpha$  and  $\Xi_{\mathbf{k}}^b$  from (III.31) as

$$\mathbf{W}_\alpha = \mathbf{F}_\alpha + \mathbf{M}_{\alpha\beta} \Delta \mathbf{R}_\beta, \quad \Xi_{\mathbf{k}}^b = \dot{\mathbf{J}}_{\mathbf{k}}^b - \mathbf{C}_{be}(\mathbf{k}) \delta \rho_{\mathbf{k}}^e \quad (\text{B.2})$$

where the summation is implied for repeated indices, and the matrix  $\mathbf{M} \equiv k_B T \mathbf{L}^{-1}$  and  $\mathbf{C}_{be}(\mathbf{k}) \equiv i\mathbf{k} J_{bd}(k) \chi_{de}^{-1}(k)$ .

We first consider the first term of (B.1)

$$\begin{aligned}\langle \mathbf{W}_\alpha \Xi_\mathbf{k}^b \rangle &= \left\langle (\mathbf{F}_\alpha + \mathbf{M}_{\alpha\beta} \Delta \mathbf{R}_\beta) (\dot{\mathbf{J}}_\mathbf{k}^b - \mathbf{C}_{be}(\mathbf{k}) \delta \rho_\mathbf{k}^e) \right\rangle \\ &= \left\langle \mathbf{F}_\alpha \dot{\mathbf{J}}_\mathbf{k}^b \right\rangle - \mathbf{C}_{be}(\mathbf{k}) \left\langle \mathbf{F}_\alpha \delta \rho_\mathbf{k}^e \right\rangle + \mathbf{M}_{\alpha\beta} \left\langle \Delta \mathbf{R}_\beta \dot{\mathbf{J}}_\mathbf{k}^b \right\rangle - \mathbf{C}_{be}(\mathbf{k}) \mathbf{M}_{\alpha\beta} \left\langle \Delta \mathbf{R}_\beta \delta \rho_\mathbf{k}^e \right\rangle\end{aligned}\quad (\text{B.3})$$

We already showed that the last two terms vanish ((III.22) and Appendix A). We now show that the two terms vanish as well. Let us first consider the second term in (B.3). We have

$$\begin{aligned}\left\langle \mathbf{F}_\alpha \delta \rho_\mathbf{k}^e \right\rangle &= \frac{1}{Z_c} \int d\Gamma_c \left( \sum_l e^{i\mathbf{k} \cdot \mathbf{r}_l} \right) \mathbf{F}_\alpha e^{-\beta \mathcal{U}} = \frac{1}{Z_c} \int d\Gamma_c \sum_l e^{i\mathbf{k} \cdot \mathbf{r}_l} \left( -\frac{\partial \mathcal{U}}{\partial \mathbf{R}_\alpha} \right) e^{-\beta \mathcal{U}} \\ &= \frac{1}{Z_c} \int d\Gamma_c \sum_l e^{i\mathbf{k} \cdot \mathbf{r}_l} \beta^{-1} \frac{\partial}{\partial \mathbf{R}_\alpha} e^{-\beta \mathcal{H}} = 0\end{aligned}\quad (\text{B.4})$$

where the last equality results upon doing the  $\mathbf{R}_\alpha$ -integration by parts. Obviously we will have the same result for the first term,  $\left\langle \mathbf{F}_\alpha \dot{\mathbf{J}}_\mathbf{k}^b \right\rangle = 0$ , since  $\dot{\mathbf{J}}_\mathbf{k}^b$  does not contain  $\mathbf{R}_\alpha$ . Therefore we showed that

$$\langle \mathbf{W}_\alpha \Xi_\mathbf{k}^b \rangle = 0. \quad (\text{B.5})$$

Next we consider the second term in (B.1):  $\left\langle \mathbf{W}_\alpha t \mathcal{Q} i \mathcal{L} \Xi_\mathbf{k}^b \right\rangle = \left\langle \mathbf{W}_\alpha t (1 - \mathcal{P}) \dot{\Xi}_\mathbf{k}^b \right\rangle$ . We need to compute the term  $\mathcal{P} \dot{\Xi}_\mathbf{k}^b$ . From (II.3), one can obtain

$$\begin{aligned}\mathcal{P} \dot{\Xi}_\mathbf{k}^b &= \frac{1}{k_B T} \mathbf{M}_{\alpha\beta} \left\langle \Delta \mathbf{R}_\alpha \dot{\Xi}_\mathbf{k}^b \right\rangle \Delta \mathbf{R}_\beta + \frac{1}{M_\alpha k_B T} \left\langle \mathbf{P}_\alpha \dot{\Xi}_\mathbf{k}^b \right\rangle \mathbf{P}_\alpha \\ &\quad + N^{-1} \chi_{ad}^{-1}(\mathbf{k}) \left\langle \delta \rho_{-\mathbf{k}}^a \dot{\Xi}_\mathbf{k}^b \right\rangle \delta \rho_\mathbf{k}^d + N^{-1} J_{ad}(\mathbf{k}) \left\langle \mathbf{J}_{-\mathbf{k}}^a \dot{\Xi}_\mathbf{k}^b \right\rangle \mathbf{J}_\mathbf{k}^d \\ &= N^{-1} \chi_{ad}^{-1}(\mathbf{k}) \left\langle \delta \rho_{-\mathbf{k}}^a \dot{\Xi}_\mathbf{k}^b \right\rangle \delta \rho_\mathbf{k}^d + N^{-1} J_{ad}(\mathbf{k}) \left\langle \mathbf{J}_{-\mathbf{k}}^a \dot{\Xi}_\mathbf{k}^b \right\rangle \mathbf{J}_\mathbf{k}^d\end{aligned}\quad (\text{B.6})$$

where the last equality holds since  $\left\langle \mathbf{R}_\alpha \dot{\Xi}_\mathbf{k}^b \right\rangle = -\left\langle \mathbf{P}_\alpha \Xi_\mathbf{k}^b \right\rangle = 0$  and  $\left\langle \mathbf{P}_\alpha \dot{\Xi}_\mathbf{k}^b \right\rangle = 0$ . Using (B.6), we have

$$\mathcal{Q} \dot{\Xi}_\mathbf{k}^b = (1 - \mathcal{P}) \dot{\Xi}_\mathbf{k}^b = \dot{\Xi}_\mathbf{k}^b - N^{-1} \chi_{ad}^{-1}(\mathbf{k}) \left\langle \delta \rho_{-\mathbf{k}}^a \dot{\Xi}_\mathbf{k}^b \right\rangle \delta \rho_\mathbf{k}^d - N^{-1} J_{ad}(\mathbf{k}) \left\langle \mathbf{J}_{-\mathbf{k}}^a \dot{\Xi}_\mathbf{k}^b \right\rangle \mathbf{J}_\mathbf{k}^d \quad (\text{B.7})$$

It is important to note that  $\mathcal{Q} \dot{\mathcal{R}}_\mathbf{k}$  only involves the solvent coordinates and momenta. Then, using the orthogonality relations  $\left\langle \mathbf{W}_\alpha \delta \rho_\mathbf{k}^d \right\rangle = 0$  and  $\left\langle \mathbf{W}_\alpha \mathbf{J}_\mathbf{k}^d \right\rangle = 0$ , one obtains

$$\begin{aligned}\left\langle \mathbf{W}_\alpha \mathcal{Q} i \mathcal{L} \Xi_\mathbf{k}^b \right\rangle &= \left\langle \mathbf{W}_\alpha \mathcal{Q} \dot{\Xi}_\mathbf{k}^b \right\rangle = \left\langle \mathbf{W}_\alpha \dot{\Xi}_\mathbf{k}^b \right\rangle = \left\langle (\mathbf{F}_\alpha + \mathbf{M}_{\alpha\beta} \Delta \mathbf{R}_\beta) \dot{\Xi}_\mathbf{k}^b \right\rangle \\ &= \left\langle \mathbf{F}_\alpha \dot{\Xi}_\mathbf{k}^b \right\rangle + \mathbf{M}_{\alpha\beta} \left\langle \Delta \mathbf{R}_\beta \dot{\Xi}_\mathbf{k}^b \right\rangle = 0\end{aligned}\quad (\text{B.8})$$

where in the last line the first term vanish from the argument shown in (B.4), and we already showed the second term vanishes.

It is now clear that the repeated applications of  $\mathcal{Q} i \mathcal{L}$  on  $\Xi_\mathbf{k}^b$  will never generate the solute-variable components, and hence  $\left\langle \mathbf{W}_\alpha (\mathcal{Q} i \mathcal{L})^n \Xi_\mathbf{k}^b \right\rangle = 0$ . Therefore we obtain the final result

$$\langle \mathbf{W}_\alpha e^{t \mathcal{Q} i \mathcal{L}} \Xi_\mathbf{k}^b \rangle = 0. \quad (\text{B.9})$$

---

- [1] K. Henzler-Widman and D. Kern, *Nature* **450**, 964 (2007).
- [2] L. Makowski, D. J. Rodi, S. Mandava, D. Hinh, D. B. Gore, and R. F. Fischetti, *J. Mol. Biol.* **375**, 529 (2008).
- [3] H. S. Mchaourab, K. J. Oh, C. J. Fang, and W. L. Hubbell, *Biochemistry*, **36**, 307 (1997).
- [4] T. P. Burghardt, M. P. Josephson, and K. Ajtai, *Biochemistry*, **50**, 7809 (2011).
- [5] W. W. Smith, C. F. Schreck, N. Hashem, S. Soltani, A. Nath, E. Rhoades, and Corey S. O'Hern, *Phys. Rev. E* **86**, 041910 (2012).
- [6] M. Feig (ed), *Modeling Solvent Environment: Applications to Simulations of Biomolecules*, Wiley-VCH Verlag, 2010, and references therein.
- [7] M. Ikeguchi, J. Ueno, M. Sato, and A. Kidera, *Phys. Rev. Lett.*, **94**, 078102 (2005).
- [8] T. Noguti and N. Go, *Nature* **296**, 776 (1982).
- [9] A. Kitao, F. Hirata, and N. Go, *Chem. Phys.* **158**, 447 (1991).
- [10] H. Mori, *Phys. Rev.* **112**, 1829 (1958); *Prog. Theor. Phys.* **28**, 763 (1962); **33**, 423 (1965); **34**, 399 (1965).
- [11] F. Hirata ed., *Molecular Theory of Solvation*, Kluwers, 2003.
- [12] T. Imai, R. Hiraoka, A. Kovalenko, F. Hirata, *J. Am. Chem. Soc.* **127**, 15334 (2005); N. Yoshida, T. Imai, S. Phongphanphanee, A. Kovalenko, and F. Hirata, *J. Phys. Chem. B*, **113**, 873 (2009).
- [13] B. Kim, S. -H. Chong, R. Ishizuka, and F. Hirata, *Cond. Matt. Phys.* **11**, 179 (2008).
- [14] R. Zwanzig, *Phys. Rev.* **124**, 983 (1961); R. Zwanzig, *Nonequilibrium Statistical Mechanics*, Oxford University Press, 2001.
- [15] U. Balucani and M. Zoppi, *Dynamics of the Liquid State*, Oxford University Press, 1994.
- [16] G. F. Mazenko, *Nonequilibrium Statistical Mechanics*, Wiley-VCH Verlag, 2006.
- [17] D. R. Reichman and P. Charbonneau, *J. Stat. Mech* P05013 (2005).
- [18] T. Yamaguchi, S.-H. Chong, and F. Hirata, *J. Mol. Liq.* **112**, 117 (2004).
- [19] F. Hirata, *J. Chem. Phys.* **96**, 4619 (1992).
- [20] The rigorous expression for  $J_{ab}$  has been derived by the authors of Ref. [18].
- [21] F. Hirata, T. Munakata, F. Raineri, H.L. Friedman, *J. Mol. Liquids* **65/66**, 15 (1995); T. Ishida, F. Hirata, S. Kato, *J. Chem. Phys.* **110**, 11423 (1999); K. Nishiyama, F. Hirata, T. Okada, *Chem. Phys. Lett.* **330**, 125 (2000).
- [22] S. Phongphanphanee, T. Rungrommongkol, N. Yoshida, S. Hannongbua, and F. Hirata, *J. Am. Chem. Soc.* **132**, 9782 (2010).
- [23] M. C. Wang and G. E. Uhlenbeck, *Rev. Mod. Phys.* **17**, 323 (1945).
- [24] G. Lamm and A. Szabo, *J. Chem. Phys.* **85**, 7334 (1986).
- [25] J. Kottalam and D. A. Case, *Biopolymers* **29**, 1409 (1990); D. A. Case, *Curr. Opin. in Struc. Biology* **4**, 285 (1994).

- [26] A. Ansari, J. Chem. Phys. **110**, 1774 (1999).
- [27] Chia C. Shih and S. Georghiou, J. Biom. Struc. Dyn. **17**, 921 (2000); Biopolymers, **81**, 450 (2006).
- [28] M. Zacharias, Biopolymers **54**, 547 (2000).
- [29] Jianpeng Ma, Structure **13**, 373 (2005).
- [30] M. Kinoshita, Y. Okamoto, and F. Hirata, J. Am. Chem. Soc. **120**, 1855 (1998).
- [31] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics*, Oxford University Press, 1986.
- [32] R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957).